DOCUMENT RESUME

ED 442 832 TM 031 252

AUTHOR Yung, Benny Hin-Wai
TITLE Biology Teachers' Struggle with the Issue of Fairness in a School-Based Assessment Scheme.
PUB DATE 2000-05-00
NOTE 23p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (73rd, New Orleans, LA, April 28-May 1, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Biology; *Equal Education; Evaluation Methods; Foreign Countries; Laboratory Procedures; *Science Teachers; Secondary Education; *Student Evaluation
IDENTIFIERS Discourse; *Fairness; Hong Kong

ABSTRACT

This interpretive study of the implementation of a school-based assessment scheme of biology practical work in Hong Kong examines the struggles of three teachers with the issue of fairness in relation to their classroom actions. Though the teachers' discourses were dominated by, and their classroom actions were pre-eminently influenced by, the notion of fairness, they did so in three qualitatively different ways: (1) fair in the sense of assessing students on a fair basis; (2) fair in the sense of not jeopardizing students' chances to learn the subject matter while they are being assessed; and (3) fair in the sense of not depriving students' opportunities of receiving an all-round education. The implication is that assessment innovation is a necessary but not sufficient mechanism for changes within the educational system. It should be undertaken with full regard to problems of teacher interpretation and mediation at the classroom level. Otherwise, this will ruin the best intentions behind the assessment innovation. (Contains 1 table and 40 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 442 832

Paper Presented at

2000 Annual Meeting of

National Association for Research in Science Teaching

New Orleans, LA, USA

April 28 – May 1, 2000

Biology Teachers' Struggle With the Issue of Fairness in a School-based Assessment Scheme

Benny Hin-Wai Yung
Department of Curriculum Studies, The University of Hong Kong,
Pokfulam Road, Hong Kong
Email: hwyung@hkucc.hku.hk

ABSTRACT

This interpretive study of the implementation of a school-based assessment scheme of biology practical work in Hong Kong examines the teachers' struggle with the issue of fairness in relation to their classroom actions. Though the teachers' discourses were dominated by, and their classroom actions were pre-eminently influenced by, the notion of fairness, they did so in three qualitatively different ways: (1) fair in the sense of assessing students on a fair basis, (2) fair in the sense of not jeopardizing students' chances to learn the subject matter while they are being assessed, (3) fair in the sense of not depriving students' opportunities of receiving all-round education. The implication is that assessment innovation is a necessary, but not a sufficient mechanism for changes within our educational system. It should be undertaken with full regard to problems of teacher interpretation and mediation at the classroom level. Otherwise, this will ruin the best intentions behind the assessment innovation.

TM031252

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

B. H-W Yung

2

BEST COPY AVAILABLE

Introduction

New approaches to assessment have emerged in a number of countries over the last several years, out of a variety of overlapping debates about the purposes and methods of assessment, and about its impact on the process of teaching and learning. Some of these debates are taking place in the context of intense political debate about the accountability of education services as in U.K. (Torrance 1995). Others, like Hong Kong, are more concerned with how the changes in assessment practices and procedures can contribute to improve teaching and learning (ROPES 1998). It was against such a background that the former Hong Kong Advanced Level Biology Practical Examination was replaced by a school-based continuous assessment scheme - the Teacher Assessment Scheme (TAS) in 1993. Advocates of the TAS claimed that its implementation would have a liberating influence on the curriculum and would bring about a host of desirable curricular and pedagogical changes.

Hargreaves (1989) aptly pointed out that assessment reforms would bring no automatic pedagogical shifts, only the creation of opportunity for alternatives to be explored more thoroughly. Findings from different studies on the effect of assessment reforms on classroom teaching were inconsistent and were somewhat context-specific. In some cases, teachers were encouraged to use a wider variety of teaching methods (Buchan 1993). In other cases, however, teachers were very much bothered by having to assume a dual role of an assessor and a teacher (Donnelly, Buchan, Jenkins and Welford 1994). In the worst situations, as observed by Buchan (1993), the formative functions of assessment were supplanted by the summative function. Teachers became reluctant to help students. Teaching and assessment remained largely polarized. Assessment did not arise as a natural consequence of teaching, nor had the outcomes been used to inform the process of teaching.

Teacher beliefs and their actions

Like many studies of change that have focused on explaining why innovations are not implemented as their developers anticipated, the above studies give only glimpses of the underlying concerns of teachers which appear to have influenced their use of the new ideas. They stop short of probing further into the teachers' thinking and developing an understanding of the change process and factors that influence the process from the perspective of the teachers. More recently, more studies have begun to explore the relationships between the beliefs of teachers and the enacted science curriculum (e.g. Briscoe 1993, Cronin-Jones and Shaw 1992, Fischler 1994, Kagan and Tappins 1991, Lakin and Wellington 1994, Tobin and LaMaster 1995). There is a growing body of evidence supporting the premise that teachers do have theories and belief systems which play an important part in their cognition and behaviour in teaching.

In studies of how teachers implemented new initiatives in the curriculum (e.g. McRobbie and Tobin 1995, Mitchener and Anderson 1989, Tobin and LaMaster 1995), it is found that when the philosophy of the curricular innovation is significantly different from the theories or beliefs the teacher holds, the challenge or demand on the teacher requires him to re-structure his beliefs or to 'domesticate' the curriculum in order to fit into his belief system. Sometimes, such re-structuring or domestication may be problematic or even 'personally threatening'. Is this the case for teachers in the TAS? TAS demands a re-formulation of their role as both as an assessor and as a teacher. Underpinning the TAS is also a radical shift in the thinking of what assessment is about, its role and its relationship to learning, so much so that these entail a shift in thinking which is referred to by Gipps (1994) as a paradigm shift - from a "testing" model to a broader model of "educational assessment". Whether these new beliefs can be translated into pedagogical form would depend on

the personal beliefs of teachers to reorganize the changes needed as well as their preparedness to meet the challenges.

Some studies (e.g. Brickhouse and Bodner 1992, Fischler 1994) have begun to look into the mechanisms teachers create to circumvent their constraints or how they adjust their teaching to fit their circumstances. Various labels have been assigned to these decision-making and/or coping mechanisms such as functional paradigm (Lantz and Kass 1987), models of reality (Duschl and Wright 1989), theory-in-use (Benson 1989), teacher mind frames (Tobin, Kahle and Fraser 1990), cultural myths (Tobin and McRobbie 1996) and an individual's awareness of the world (Marton 1994). In the present study, as the data emerged, it became clear that Marton's phenomenological approach to the problem was a useful conceptual tool to make sense of the data.

According to Marton and Booth (1997), a person's awareness is "the world as experienced by the person" (p. 108). That is, awareness is the totality of all experiences of an individual in terms of which a certain phenomenon is understood; an experience being an internal relationship between the person and the world. They also argue that "there is only a limited number of distinctly different ways in which people are capable of experiencing any of the things they meet" (p.206). And that the variation in the different ways that people experience something can be understood in terms of the limited human capacity to discern and to be aware of diverse aspects of situations and phenomena at the same time. In short, a particular way of experiencing something represents a combination of related aspects that are simultaneously present in a person's focal awareness. In the present study, the phenomenon is the notion of fairness as experienced by teachers in the context (or 'situation' - in Marton's term) of a school-based assessment.

According to Marton, there are various aspects of teaching, of aims of education in general, of aims of science education, of aims of practical work instruction, of the particular content and of the assessment requirements, in relation to which teachers differ significantly. These aspects are different layers of the teachers' awareness, which are simultaneously present. In teaching a certain content, all the increasingly general aspects of the teachers' awareness including those of the cultural myths are inherent. Everything is there - admittedly not always clearly; not always explicitly - in every moment of the very teaching of the very content. They would only be drawn into consideration when circumstances arise. Among the various general aspects of the teachers' awareness, "pressures of accountability" is commonly reported by many of the studies (e.g. Duschl and Wright 1989, Brickhouse and Bodner 1992) as a constraint affecting teachers' pedagogical decisions. Given Hong Kong's situation, where the emphasis on examination for selection purposes still is much stronger than in other places (Biggs 1996, Morris 1985), it is anticipated that this and other related concerns like "to help students pass the examination" and "to maintain student respect" etc. are bound to be among local teachers' structure of awareness. Is this really the case? And this leads us to the focus of the present study.

This paper reports on the classroom assessment practices of three biology teachers in the context of the TAS. It examines their pedagogical decisions and associated beliefs which informed their decisions. One emergent theme from the data was the teachers' concern of the notion of fairness. The three teachers had been trying to make their classroom practices as fair as possible to all their students while carrying out the assessment. They were all struggling with the issue of fairness in implementing the TAS but in qualitatively different ways. The paper first discusses the perceived educational merits of the TAS so as to provide readers with a background on what the expected outcomes were. It then reports on the classroom practices of the three teachers and their associated thinking and beliefs. The consequences and implications of these classroom practices are also discussed.

BEST COPY AVAILABLE

Perceived educational merits of the HKAL Biology TAS

The HKAL Biology TAS requires teachers to assess their own students on five aspects - practical skills, report-writing, dissection, field work as well as attitude towards practical work. Prior to its implementation, it was argued that the TAS would bring about several educational benefits (Pang 1992). Namely, the TAS can provide a more valid and reliable assessment of students' practical skills as well as a wider measurement of students' achievement of educational objectives, in particular, students' affective characteristics. It can motivate students because only the best marks from a number of practical tasks are taken for each student. TAS also helps teachers in providing more feedback to students. It can have a liberating influence on teaching. Since there is no further need of drilling for particular types of practical work required by practical examinations, more varied, meaningful and relevant experiments and projects can be conducted. This will allow better integration of theory with practical work. Indeed, the 21 teachers participating in a pilot scheme of the TAS prior to its implementation expressed similar views towards the educational values of the TAS (Yung 1992).

Pang (1992) also argued that implementation of the TAS would provide opportunities for the professional development of teachers via various in-service training activities. In a study conducted 2 years after the implementation of the TAS (Yung 1995), most teachers (N=222, representing 67% of the total population) confirmed that the TAS had contributed significantly to their professional development. On the knowledge side, teachers felt that they had come to realise the importance of assessing both the product and process of students' practical work, the importance of assessing students' affective characteristics, the importance of integrating teaching with assessment. They thought that they had also gained a better insight into the role of practical work in biology teaching and a better understanding of the functions of formative assessment and those of summative assessment. With regard to skills, the teachers reported that the TAS had contributed significantly to their skills in assessment and to their ability of inculcating in students the appropriate affective qualities. In addition, teachers felt that they had also gained a better access to useful ideas for teaching and assessment. Overall, teachers acknowledged the contribution of TAS to their professional development; and thought that they had become more accountable to students and professional colleagues.

However, it should be noted that the two studies cited above (Yung 1992, 1995) were quantitative studies where teachers were asked to express their views on *a priori* statements using a Likert scale. Such a reductionist approach to the problem did not tell much about what was actually taking place inside the classrooms. Indeed, the naturalistic qualitative study reported in the next section uncovered quite a different story.

Implementational problems of the HKAL Biology TAS

Data sources

The study reported in this paper is part of a bigger study comprising more than 40 classroom observations of 10 biology teachers. The lessons were recorded via a wireless microphone attached to the clothing of the teacher. The observer sat at a far back corner of the classroom monitoring the recording as well as jotting notes of the events taking place. The classroom observations were supplemented by post-lesson interviews which probed teachers' pedagogical decisions as well as their associated thinking and beliefs. Numerous interviews were conducted with the participants. These include a biographic interview, teacher's narration of a good quality practical work and a good practical work assessment, Conception of Teaching Science Interview (Hewson and Hewson 1989). However, not all these data sources were employed directly in this study, although their use in other studies provides in-depth insights into the context in which the present study is embedded.

Data treatment

Interpretations of the data collected were constructed using a method similar to that described by Erickson (1986). Data were analyzed on a continuous basis throughout the study. The analysis consisted of reviewing and comparing the data sources, and looking for patterns and irregularities. Interpretive research is described by Erickson as focusing on "the immediate and local meanings of actions, as defined from the actors' point of view" (p.119). Accordingly, the prime aim of this study was to find out, from the perspectives of the actors, the 3 biology teachers, what was happening in their classrooms and why they acted in a certain way.

Procedures were undertaken to ensure that the study would yield outcomes that were authentic, trustworthy, credible and robust. For example, the use of numerous data sources maximised the probability that the emergent assertions were consistent with a variety of data. Any evidence that was counter to an assertion was explored in detail, and care was taken to build an understanding of the discrepancy. In this process, assertions were elaborated so as to take account of all data. Serious attempts to refute assertions added to the credibility of the findings of the study. Member checks (Guba and Lincoln 1989) were also used for the same purpose. Data and interpretations were shared with participants so that they can review the data and interpretations, agree or disagree with the assertions of the research, and suggest corrections, elaborations and summary statements.

Viewing the research from a constructivist perspective, the knowledge claims presented in this paper cannot be viewed as representing a necessarily true or correct portrait of the factors that influenced the teachers as they implemented the TAS. Rather, they represent viable explanations which emerged from the data. The extent of transferability of the findings of this study to contexts that readers might encountered will be constrained by the perspective of the reader and the potential applications the reader might consider. That is, the readers have to decide for themselves what are the potential applications, if any, of the findings to their own context. For this reason, it is important for me to delimit the extent to which findings from this study can be applied through a detailed description for each of the case studies so as to provide sufficient information for readers to assess the potential transferability, appropriateness for their own settings (Guba and Lincoln 1989). To this end, I decided to use narrative accounts as the way to convey the context of this study and the knowledge that was implicit in the stories of the 3 teachers and of myself who was a participant observer. In this way, I felt I could emphasis the voices of the teachers in a credible manner, and emphasise my interpretations in the latter parts of the paper. Pseudonyms are used for the teachers in the following narrative accounts and interpretations.

Narrative accounts and interpretations

Reported below are classroom episodes in 3 of the lessons observed and the associated thinking of the teachers (A, B and C) concerned, focusing on how the teachers made sense of assessment as they implemented the TAS. The 3 lessons were selected according to two criteria. First, the same subject content was being taught and/or assessed. Second, there were many variations among the teachers' actions in the classroom.

In each of the 3 cases, students were asked by their teacher to carry out an investigation related to water relation in potato cells. They had been taught the necessary biological theory but they did not know the exact problem to be solved. They came to know of it only at the beginning of the lesson. After this they were not allowed to refer to textbooks or any other references. At the end of the experiment, they handed in a report for assessment. Teacher A and B, but not C, also assessed students on their experimental skills during the course of the experiment.

Figure 1 provides some background information regarding the teachers and the students. It also provides an overview of the lessons in terms of the duration of the different phases of the lessons (prelab discussion, lab work, post-lab discussion, report writing) and the amount of teacher-student interaction. One dialogical text unit refers to the part of a dialogue exchange [between the teacher and the student(s)] which is followed or separated by a natural break or pause in speech. That means, the number of dialogical text unit gives a rough indication of the relative amount of interaction between a teacher and his students during the practical. That is to say, Teacher A interacted far less with her students (124 dialogical text units) when compared with Teacher B (904 dialogical text units) and Teacher C (692 dialogical text units) during the practical observed. This is somehow related to their perception of their role as a teacher in general and in the TAS in particular. Relevant episodes of the teachers' actions and the rationale they offered for the actions are highlighted below to illustrate how the teachers' thinking could have influenced their classroom teaching, in particular, their readiness in helping the students to overcome their learning difficulties.

[*Insert figure 1 here]

Teacher A - "I must be fair."

Mrs Chan was in her late forties and has taught AL Biology for more than 16 years. She was teaching a class of 22 girls in a prestigious school. She informed students in advance that this practical was going to be an assessment practical. That is, special assessment sessions were set rather than such skills being assessed more realistically and authentically in the course of ordinary practicals as originally intended by the TAS.

Episode 1

At the students' request, Mrs. Chan allowed students to come into the laboratory 15 minutes (designated as the 'preview' period in Figure 1) before the practical was officially due to start.

The rationale behind Mrs. Chan's decision to allow students to come in earlier was:

> "They (the students) take assessment very seriously. They want to have a look at the apparatus first and sort of prepare themselves psychologically for the practical. So, I allow them to come in earlier during the lunch hours ... Some of them have very high expectation of themselves ... *They behave this way whenever marks are counted,* be it theory or practical assessment. *They have been brought up this way ...*" (my emphasis)

Students' obsessive concern with assessment was a prominent feature in Mrs. Chan's classroom as in many other classrooms in Hong Kong. This is indeed a culture as Mrs. Chan has put it – "they have been brought up this way". The act of allowing students to come into the laboratory earlier in itself was also heavily laden with assumptions rooted in cultural myths – "taking assessment seriously … high expectations … whenever marks are counted," etc. Thus, instead of alleviating the pressure on students associated with the one-off external practical examination, TAS – in the way it was implemented by Mrs. Chan - had in effect aggravated the situation by exerting a lot more pressure onto students continually throughout the 2-year A-level course. The influence of student's anxiety about the consequences of success or failure as shown in the above episode must not be overlooked. Though Mrs. Chan had rightly put the blame to the culture, she might not be aware that she herself was amongst one of those who were perpetuating this culture unknowingly as could be seen in the following episode.

BEST COPY AVAILABLE

Episode 2

Mrs. Chan started the lesson by distributing the laboratory manual to the class and allowed them some time to read it. She then invited questions:

> "This is a Type I experiment (where students are expected to be given very limited guidance and to work on their own) ... Any questions before we start? Any questions, please?"

There were no questions from the students. Mrs. Chan then said again,

> "Come on, any question? *Free of charge! Marks will not be deducted.* Come on. Any question?" (my emphasis)

Again, there was no question from the students. So Mrs. Chan signalled the class to begin their work.

As suggested by Marton (1994), learning and teaching (as well as assessment in this case) take place in the social space of the classroom. It is there in the world to which the phenomenon belongs. It is shared, it is an object of awareness. It is not just a mentalistic thought object residing inside the solitary head of the solitary teacher. Why was there no question from the students? This was probably because the amount of help provided to students constituted one of Mrs. Chan's criteria for assessing her students' practical competence. Marks would be deducted from the student for the help offered. Students in general did not like the idea. Often they preferred to proceed without assistance even though they realized that they might not be able to generate an effective response to the problem assigned. Equally Mrs. Chan was also reluctant to answer students' questions during the course of the experiment as shown in the following episode.

Episode 3

In the middle of the practical, a student wanted to ask Mrs Chan a question. The dialogue was as follows: (S is the student and T is the teacher)

S: I have a question but would marks be deducted?
T: You ask it first.
S: Chee! I don't want to ask then.
T: If I am going to deduct marks, I will tell you first.
S: If I ask you the question, but then you tell me afterwards that marks have been deducted, I would be very depressed.
T: Just go ahead and ask me. And you will know what the outcome would be.
[The student then asked the question.]
T: I have to deduct marks from you if I answer you this question. Therefore, I am not going to answer this question. You think about it yourself.
S: Are you really not going to deduct any mark from me at all?
T: Go back and do your work quickly.

The reason behind Mrs. Chan's decision not to answer the student's question was that:

> "*I must be fair*. I can't answer some students' questions but not the others. If I answer her question, I am sure that I will have to tell the rest of the class using the microphone ... What *bothers me* is that suppose I am going to answer students' question, how many questions should I entertain, and to what extent? This is the *most difficult* part ... *If there is no TAS, I would then give her a definite answer* ... " (my emphasis)

Mrs. Chan was caught in a dilemma of trying to be fair to all students on one hand and trying to help solve their problems on the other. Obviously, teaching and assessment became polarized in the above situation; and teaching had given way to assessment. The formative function of assessment was lost. Mrs. Chan's perception of assessment as purely for summative purpose tempted her to undertake assessment in formal testing situations where students' requests for help were not entertained. In the long run, such a practice could have an adverse influence on the student-teacher relationship as implied in the following two episodes.

Episode 4

Mrs. Chan was standing next to a student watching her doing her work, the student said to her, though in a joking manner,

"You go away because you didn't answer my question."

When asked to express her feeling on this instance, Mrs Chan said,

> "*She won't blame me* just because I didn't answer her question. I think that she just did not want me to disturb her." (my emphasis)

One of the important external influences on learning is the classroom environment, including the student-teacher relationship, and the teacher's style and competence, as these determine the overall quality of the classroom climate the teacher creates and the quality of her interaction with individual students. In the above instance, Mrs. Chan might not be aware of the possible damaging effect of not entertaining students' questions on the student-teacher relationship. The possible deteriorating student-teacher relationship could be revealed in the following episode where it took Mrs. Chan some effort (extending her help by asking a student four times before getting a positive response) to get to know what sort of difficulties a student was encountering.

Episode 5

Mrs. Chan noticed that two pupils were discussing. She had a feeling that they might be encountering some difficulties about the practical. Thus she intervened and asked one of the girls:

T: Is that you don't know how to do it?
S: No, no problem. No problem.
T: Any problem?
S: No problem. No problem.
T: Just ask me if you've got any problem.
S: No problem.
T: Is there any problem? You don't know how to do it? Any problem?
S: I forget what is the meaning of 'hypotonic'.
T: Hypo means a greater water potential.

When asked about her feelings about this instance, Mrs. Chan expressed her feeling of being disempowered by the TAS though rather implicitly:

> "Students generally do not want to reveal their difficulties or weakness to the teacher. Maybe their impression is that marks are going to be deducted ... I just don't know why she did not want to ask me. *What can I do?* " (my emphasis)

Obviously, the effect of TAS on the teachers' professional situation was not so uniformly positive as anticipated. This episode illustrates how Mrs. Chan's professional situation was undermined.

Before the implementation of the TAS, teachers were broadly free to run their class and laboratories as they saw fit. There were accepted canons of practice and a body of "teacherly" knowledge about how things were done. The latter is a form of tacit and practical knowledge which is sustained through traditions (Nyiri 1988). However, the introduction of the TAS requires teachers to construct a novel professional practice which entails a shift in their understanding of the role of assessment from within an 'educational assessment paradigm' instead of the traditional 'testing paradigm'. However, Mrs. Chan's adherence to the traditional "testing paradigm", as evident in the episode below, has made her felt somewhat dis-empowered.

Episode 6

Mrs. Chan found that most students were stuck at one point during the practical. She was talking to herself fearing that none of the students would be able to accomplish the task. Instead of circulating supervising students' work, she went back to the teacher bench and sat there with her hands supporting her head, engaging herself in deep thoughts for at least 2 minutes before announcing to the class a crucial step in carrying out the practical.

In the post-lesson interview, Mrs. Chan admitted that she was struggling in her mind if she should give the hint to the class or not. She explained:

> "I had never thought of telling them everything (about the procedure of the experiment). Given the abilities of my students, if you tell them the crucial point, you are just *making troubles to yourself*. It will be difficult for me to *differentiate* them (rank order their performance ..." (my emphasis)

What do the episodes tell us?

These episodes collectively make up an instructive case study of how firmly held ideas about assessment can affect the quality of teaching and learning at the classroom level. Mrs. Chan's first and foremost consideration were differentiation of the students' capabilities and the notion of fairness rather than wider concerns with how assessment could best be used to promote learning. The particular tensions and pressures being perceived by Mrs. Chan can be further revealed by an excerpt from the interview:

> "*TAS is certainly an assessment*. I have been criticised by my chemistry colleague that TAS is a training and not an assessment. But, *the hard fact is that I have to submit the marks to HKEA .... It is no good emphasising too much on learning* ... If you tell students that marks are not important, just relax and try. Students just would not believe in what you say especially students in this school. They know what is going on. *You just can't fool them.*" (my emphasis)

The requirement of submitting the TAS marks to HKEA for certificating purposes had framed the way in which the assessment innovation was interpreted by Mrs. Chan. She had drawn on her previous experience and understanding of what assessment was about in order to make sense of the changes and make decisions about how she should implement the TAS. The above episodes tell vividly how such an interpretative framework was deeply seated in Mrs. Chan's mind. Hence, she felt under pressure to set aside special assessment occasions that were more controlled than routine practical work sessions and not to answer students' questions during the course of the practicals. This is not surprising, in particular, if the introduction of school-based assessment was

BEST COPY AVAILABLE

thought to be purely for the purpose of improving the validity and reliability of the assessment. In fact, this is not an unreasonable assumption for teachers to make when the innovation was initiated by the HKEA – the public examining body. No wonder, with such a mindset, the classroom situation in which practical work occurred was that of a formal assessment. Though this did not mean that the teacher could not help the students, such assistance was integral to the assessment process, resulting in the deduction of marks. That means, the teacher was explicitly acting under the authority of the HKEA, as an extension of the examination procedures. This aspect of work was a source of tension for many of the biology teachers including Mrs. Chan. For these teachers, the concern of how the innovation may broaden the curriculum and improve the quality of teaching and learning will subside quietly into the background.

In summary, this case study shows that genuine improvements in the effectiveness of learning actually requires a major rethink in the way that assessment is used. This rethink needs to be based on a careful analysis of how and why assessment can promote individual learning. In particular, in-service training for teachers aiming at re-constructing their ideas of the role of assessment in the educational reform and helping them to develop strategies for utilising the outcomes of assessment in formative, evaluative and educative respects are imperative (Hodson 1992). To this end, the following 2 case studies are illustrative.

Teacher B - "They are learning while I am assessing them."

Mr. Hong was 28 years old. He had taught AL Biology for 6 years. He was teaching in a prestigious co-educational school. There were only 8 students in the class. When compared with Teacher A, Mr. Hong seemed to be able to cope with the requirements of the TAS quite well! He also seemed to be able to come to grips with the formative function of the assessment. He always initiated discussions with individual students (904 dialogical text units compared with Teacher A's 124 dialogical text units) during the course of the experiment. In fact, he was mindful of the importance of interacting with students even at the planning stage:

> "One of my major considerations in selecting practicals for the TAS assessment is whether I can make use of the practical to induce some kind of discussions with my pupils and that they can learn through it, something that they have not thought about it before. This is a very crucial part in their learning. That's why I always ask them questions continually throughout the practical. So, they are in fact *learning while I am assessing them.*" (my emphasis)

Salder (1989) believes that assessment is truly formative only when it involves the pupil. She conceptualises formative assessment as being concerned with how judgements about the quality of students' responses can be used to shape and improve their competence by short-circuiting the randomness and inefficiency of trial-and-error learning. This is exactly what Mr. Hong was trying to achieve when he engaged in active discussions with his students.

Another common feature in the discussions of Mr. Hong with his students was that he always responded to his students' queries with remarks and questions like:

"What do you think?"
"What better procedure can you think of?"
"You think it over yourself first. I will come back to you later."

In a post-lesson interview, Mr. Hong indicated to the researcher of the positive effect of the TAS on his teaching and the learning of his students:

> "In the past, I would point out to them their mistakes directly. Now, I have to remind myself to be conscious of this. Telling them directly is the fastest and simplest way. But it does not make them think. This is a good influence to both teaching and learning."

The above indicates that Mr. Hong was, at least, beginning to realise the importance of not only providing feedback to students but also the quality of the feedback as Salder (1998) has pointed out,

> "Formative assessment does make a difference, and it is quality, not just quantity, of feedback that merits our closest attention. By quality of feedback, we now realise we have to understand not just the technical structure of the feedback (such as its accuracy, comprehensiveness and appropriateness) but also its accessibility to the learner (as a communication), its catalytic and coaching value, and its ability to inspire confidence and hope."

Teacher B was able to find a handle or frame of reference outside the concrete situation of assessing his students. That is, "to be conscious of not telling students the answers directly so as to make them think". He saw this as a good influence to both his teaching and his students' learning. Marton (1994) describes such a transcendence of one's taken-for-granted experiential world as the ascent of a kind of analytic awareness: a capability of abstracting aspects of concrete situations and seeing these aspects of concrete situations in relation to each other. It seemed that it was exactly the lack of this analytic awareness in teacher A such that she was not aware of the possible damages of her classroom actions to the teacher-student relationship.

Even though Mr. Hong did not restrain himself from discussing with students as Teacher A did, he was still conscious of the issue of fairness. He felt at ease in discussing with his students on an individual basis because he felt that he was in control of the agenda and he was able to know what was going on during the discussion. However, he did not want students to discuss amongst themselves for too long because they would have compared notes and it would not be fair. His concern can be revealed in the following excerpt from an interview:

> "I won't intervene (students' discussion) unless they have been discussing for a long time. I think this is okay. This may create a more relaxed atmosphere..... In fact, this is a difficult problem for me. Suppose, if there is no assessment, they will learn more from each other through the discussions."

Mr. Hong's concern of fairness can be further revealed in an instance where he was sort of unwilling to provide further hints upon a student's request. He told the student,

> "I guided your classmates in the same way as I have done it for you. All of them can do it. What has happened to you?"

In the above episode, Mr. Hong was judging the student's performance on a norm-referenced basis, comparing the student performance with his classmates. According to Harlen and James (1997), formative assessments should always be made in relation to where students are in their learning. That is, formative assessment should be criterion-referenced and pupil-referenced (or isaptive). This means that assessment of a student's work should take into account the particular context of the student's work and the progress he has made over time. In consequence, the judgement of a piece of work, and what is fed back to the student, will depend on the student and not just on the relevant

criteria. The justification for this is that the individual circumstances must be taken into account if the assessment is to help learning and to encourage the learner. If formative assessment were purely criterion-referenced it would be profoundly discouraging for many students who are constantly being faced with failure. The point to be made here is that whilst norm-reference assessment might help teachers recognise the existence of a problem, it can offer no help in knowing what to do about it and may simply have a deleterious effect by labelling students. Obviously, in this particular instance, Mr. Hong's priority of trying to be fair to all students had made him lose sight of his obligation to construct a theory of effective learning which takes contextual variables into account including students' personal variables. This was unlike the case reported below where the teacher was highly conscious of the effect of contextual variables in affecting students' learning.

Teacher C - "Is it really fair?"

Mrs. Lee was 44 years old. She had taught AL Biology for 18 years. She was teaching a class of 17 students of average abilities in a co-educational school. In the lesson observed, there were a lot of discussion both between the teacher and students as well as amongst students themselves. For example:

During the prelab discussion, the teacher said to the class:

> "Yes, it is more accurate. But why is it more accurate? Even though we are running out of time, I want at this point, you spend about a minute to discuss with your classmates why using ... is more accurate? ... I will come back to you after one minute."

When asked why she often encouraged students to discuss amongst themselves and whether this would create a dilemma for her in coping with the requirements of the TAS, Mrs. Lee replied in the following manner:

> "This is a compromise to students' cultural habits of not wanting to be vocal. They are passive. They are unable to respond promptly. I have to give them time to think, to process and to discuss their ideas so as to *build up their confidence* ... I am aware of the conflict between teaching and assessment but there is no such formal statement about the Do's and Don'ts in the TAS Handbook. I think limited discussion won't affect their overall performance too much. Too assessment oriented will hinder a lot of ideas flowing out. *They have undergone the educational process. Is that really going to affect the fairness of the assessment? ... The interaction amongst themselves and between us is an unexplored treasure*. I have been encouraging them to speak up. But this has to be built up slowly step by step ... *I have faith in my pupils* ..." (my emphasis)

The following episode illustrates how Mrs. Lee tried to tap into the 'unexplored treasure' of interacting with her students during the postlab discussion:

T: How many of you have recorded the time you started the experiment? I mean the immersion time.
[A few students put up their hands.]
T: Good. I did not tell you to do that. But you have to report the immersion time in your report. Right?
T: Now, how long should the immersion time be?
S1: Half an hour.
T: Why should it be half an hour? Why not 35 minutes? What is the reason behind?
S1: It is not too long and not to short.
T: Why don't you say 20 minutes then? Is that only a subjective judgement? You don't have any reason behind, do you?

T: Theoretically speaking, how long should the immersion time be? How long should it be? Chan Ting Hong.
S2: Half day.
T: Half day? Some say 30 minutes. Some say half day. How could there be such a big difference?
T: What do we mean by an end-point?
T: Ho Chi Lit. What do we mean by an end-point? What is the end-point for this experiment?
S3: When we obtain the results.
T: When we obtain the results! [*laughing*] How are we going to judge whether we have obtained the results or not?
S4: When there is no net movement of water into and out of the cells.
T: Yes, it is the end-point when there is an equilibrium. The potato cell does not gain water or lose water.
T: But how do we know? Theoretically speaking, how do we know that it has reached the equilibrium state? Do you know how?
[No response from the class.]
T: You take it (the potato strips) out. You measure the weight. What would you do next? How can I tell if there would be further changes in its weight?
T: Take it out. Measure the weight. What should we do next?
S5: Put it back into the solution.
T: Put it back. Keep it immersed for some time. Take it out and weigh it again. When the weight remains unchanged, it has reached the equilibrium state then.
T: But for practical purpose, there are a number of disadvantages. One, we can't afford the time required (for such a tedious procedure). Two, …
T: Maybe half an hour or 45 minutes is quite okay. Especially when the purpose of the present experiment is only to find out the gradation (arranging the sucrose solutions in an order of increasing concentration). Right? Even if the end-point has not been reached, it would not affect our interpretation of the results. It is because we just want to find out the sequence of the gradation.
T: I don't know whether you understand what I am talking about. Do you understand?
T: What did I say?
S6: Since we only want to find out the relative concentration of the solutions, we just need to find out which increase in mass is the least and that is the solution with the highest concentration and so on.
T: What do you mean? Suppose all the set-ups have not reached the end-point, what would be the results?
S6: The set-up with the smallest increase in mass is the most concentrated solution.
T: Say, if it needs one hour to reach the end point but you take measurements only after 45 minutes. Yes, it is not so good. But I don't think that it is of a great problem. You may understand what I am saying. But you may not be able to express yourself. But you have to write it down in your report.
T: Let me explain again. If the change in weight in each set-up is smaller in amount, all set-ups will have a smaller change in weight by the same proportion. Okay?
T: That means, as long as we can rank order the percentage changes in weight in a sequence, we can achieve our purpose already.
T: When you write out this idea, you have to spend some effort. Okay?
T: What else have you found out that might have affected this investigation? And also things that you can avoid? *Each of you makes one suggestion in turn.*
S7: Begin with whom, sir?
T: You first, then.
S7: The variation in the time when we put the potato strips into the different solutions.
T: Okay. There are variations in the time when you put them into the solutions and also the time when you take them out. But we can do something to improve it. How?
S8: Have 8 persons to put the potato strips into the solutions at the same time.
T: 8 persons to perform one investigation. [*laughing*] The idea is correct. But how about if we got only one person to perform the investigation and that we don't have any special device with us?
T: Assuming that you put the strips into the solutions sequentially, what should you do when you take the end point measurement?
S8: Follow the same sequence.
T: Follow the same sequence. Right! This is to keep the variation in immersion time as little as possible.

T: Good. What other ways can you think of to improve the experiment?
S9: .....

[The class discussion continued for a few other minutes on several other points before the teacher rounded up the discussion in the following manner.]

T: Okay. I leave the rest (of the discussion points) for you to think it about yourself. I will leave you alone without further intervention. *But you can ask me questions, if necessary.* Basically, you have to work on your own. Okay? (my emphasis)

As could be revealed from the closing remark above, in addition to encouraging students to participate in the class discussion, Mrs. Lee also encouraged students to ask her questions if deemed necessary while they were writing up their reports. This was very much different from Teacher A who had had no interaction with her students at all in the report writing stage (see Figure 1). When asked if interacting with students so much would affect the fairness of the assessment, Mrs. Lee's view was:

> "*This is what science education is about.* TAS never prohibits teachers to respond to questions raised by students. Students' overall performance will not be affected by just *one or two points* which they might have discussed with the teacher or their classmates. Differentiation (in their capabilities) will be reflected in their *overall* performance in the reports..... *The idea of the TAS is to integrate assessment with teaching and learning.*" (my emphasis)

Very clearly, Mrs. Lee's way of implementing the TAS was very much related to her 'scaffolded' view of learning in which the teacher should try to provide a stimulating environment and guide her students towards learning. Thus, of the three teachers reported here, Mrs. Lee seemed to be the one who could best integrate assessment with teaching and learning. Nonetheless, in one of the interviews she also raised a point related to the notion of fairness. As could be revealed in Figure 1, it was a common practice for the teachers to extend their practicals beyond the normal school hours. Often, the extension could be more than one hour (see Figure 1). Mrs. Lee felt that this was unfair to the students:

> "I fear that pupils may feel bored when I ask them to stay after school every biology practical. This is impossible when we are emphasizing all-round education. Students are encouraged to participate in more extracurricular activities. *It is unfair to them* if they are denied of these options." (my emphasis)

Asking students to stay after school until they had finished writing up their reports was in keeping with the TAS regulation that the teacher has to exercise control and supervision over all work assessed. The public examining body thinks that this will ensure that the work assessed is the students' own and thus plagiarism will be prevented. No one will deny the importance of fairness in a public assessment system. But the key is how to achieve optimum fairness for the assessment's purpose while at the same time still facilitating teaching and learning. We should look at the problem not only at the level of teaching and learning in individual subjects but also in the larger context of all-round education as Mrs. Lee has put it. There is no simple answer to this problem. This is yet another area where professional judgement would be called in during the decision making process. This would in turn be influenced by the teacher's belief system as could be revealed in the reason put forward by Mrs. Lee for allowing her students, at times, to finish writing up their reports at home:

"There are two main factors that differentiate the students. First is his own background knowledge and the language ability. Second, whether he put efforts into it or not. These are the two overriding factors ... If he is not good in language ability or poor in background knowledge, he is impossible to improve very rapidly overnight. You can tell immediately if you interrogate him with a few questions. You may say that I have prejudice against some students."

It is understandable that both the examining body and teachers are concerned with improving the credibility of the assessment by making it as fair as possible for everyone. However, the tension between assessing under standardised conditions and providing flexibility to cope with contextual differences in different classrooms is always there. The important point to emphasize here is that the educational benefits derived from the TAS are at a cost to reliability. This cost is paid once we have decided to go for the TAS. This concurs with what Harlen (1994) says, "We must recognise that assessment in education is inherently inexact and it should be treated as such." Nevertheless, this is not to deny the importance of reliability (and hence fairness) because an unreliable assessment is not only of little use but can be unjust too. The endeavour to increase reliability is common to all methods of assessment but the context and purpose of assessment will affect the degree of priority given to reliability. The key is how to achieve optimum reliability for assessment purposes while maintaining high validity. That is, assessments that are arising as a natural consequence of teaching and learning, but not something add onto it. To this end, differences in individual classrooms must be taken into account.

In fact, there are many methods of moderation to address the issue of fairness in school-based assessment arising from variations in the marking standard of teachers in different schools and the conditions under which the assessments are carried out, etc. However, this is outside the scope of this paper. But certainly, any form of assessment leading to certification cannot be so low in reliability that its validity is seriously called into question. The other important thing is that we have then to find ways to maximize the educational benefits that can possibly be derived from implementing the TAS (as Teacher B and C did) instead of feeling bound by it (as Teacher A did). Otherwise, we are doing injustice to our students – depriving their opportunity to learn the subject matter and to receive all-round education.

Summary and implications

The study reported in this paper has shown how an exploration of classroom assessment practices of this kind – adopting a naturalistic and interpretive approach - can illuminate how teacher's belief system, in particular, their conceptualization of the role of assessment can inhibit or facilitate learning. It brings into light more vividly the kind of teaching-learning activities that are actually occurring inside the school laboratory and the various confounding variables that are affecting laboratory teaching and its assessment. Studies of this kind are definitely more informative than the previous ones which adopted a quantitative and reductionist approach (e.g. Thompson 1978).

The structure of teachers' awareness of the issue of fairness

Of the 3 cases reported, although the teachers' discourses were dominated by, and their classroom actions pre-eminently influenced by, the notion of fairness, they did so in three qualitatively different ways:

- as an extension of the public examination procedure (Teacher A)
- providing chances for students to learn the subject matter (Teacher B)
- providing students with an all-round education (Teacher C)

Given that the teachers had been given little support by way of a model for school-based assessment, it was not surprising to find that they adopted a range of procedures. Though the teachers were carrying out assessments under the same set of regulations governing the TAS, they did so in terms of goals of assessment, in terms of goals of learning and teaching and in terms of abilities that the students were supposed to develop, etc. These specific goals were not considered in isolation but were seen in relation to more general goals such as:

- an extension of the public examination procedure as reflected by the following excerpts from the interviews with Teacher A:

  "TAS is certainly an assessment ... It is no good emphasising too much on learning ... I have to be fair ... I have to differentiate them", etc.
- providing chances for students to learn the subject matter as reflected in the following statements made by Teacher B:

  "They are in fact learning while I am assessing them ... [I want] to make them think."
- providing students with an all-round education as reflected in the following statement made by Teacher C:

  "We are emphasizing all-round edcuation. Students are encouraged to participate in more extracurricular activities. It is unfair to them if they are denied of these options."

In turn, the specific goals were also seen in relation to the ways in which the teacher could possibly contribute to bringing about those more general goals. For example:

- restraining from giving clues to students to help them solve their problems or answering students' questions and not allowing students to discuss among themselves, etc. were ways considered by Teacher A to fulfill her goal of acting as an extension of the public examination procedure.
- asking students a lot of questions so as to make them think was a way considered by Teacher B to be effective in providing chances for students to learn the subject matter while he was assessing them.
- allowing students to complete the laboratory reports at home so that they would not be deprived of the opportunities to participate in extracurricular activities after school was seen by Teacher C as contributing to her goal of providing students with an all-round education.

Accordingly, dealing with goals of learning in a specific sense implied thus in the above cases – explicitly or implicitly – dealing with relations. Relations between content of the assessment task and student, relations between specific goals and general goals, relations between goals and means. All these depended on the teacher's intentionality; his directedness, what he was oriented towards and in which way (Marton 1994).

Thus, Teacher A's consciousness was directed to maintaining fairness in the differentiation of students' abilities. She saw herself as acting under the authority of the HKEA, as an extension of the examination procedures. The concern of how formative assessment could improve the quality of students' learning had subsided quietly into the background. Certainly, this was not to the best interest of the students concerned and was grossly unfair to them as they were denied of the opportunities to learn when the teacher refused to entertain their questions.

Teacher B, for most of the time, was conscious of making the assessment as truly formative as possible by intentionally involving pupils in discussions with him. He wanted students to learn too while he was assessing them. He was directed to this goal even in the planning stage of which practical was to be used for the assessment. This was, of course, a much fairer practice to the

students when compared with that of teacher A in terms of providing students with the opportunity to learn the subject matter.

On the other hand, though teacher C could integrate assessment with teaching and learning of the subject matter very well, she still found the present assessment practice unfair to the students when she looked at it from the perspective of providing students with an all-round education. This was because, she thought, asking students to stay after school until they had finished writing up their reports was denying their opportunities to participate in more extracurricular activities which she saw was contrary to her goal of providing students with an all-round education.

Teacher C was actually discerning relevant aspects in the situation and relating them to each other in functional terms. The relations were, in this case, between school-based assessment and students' learning, and between school-based assessment and all-round education for students. This latter consideration brings out a salient point which is often overlooked by public examining bodies in the process of striving for a credible school-based assessment. That is, the assessment should not make demands on teachers and students that are incompatible with the context in which learning is exhibited. In short, the assessment should not be too time consuming, artificial nor divorced from the normal range of contexts in which the educational achievements of students can be observed. Clearly, the extension of the practical sessions beyond normal school hours in all the three cases reported above is a case in point. In so doing, this has put the validity and the educational desirability of the TAS seriously in doubt.

Developing teachers' analytic awareness on the issue of fairness

Many of the problems revealed are, in fact, contentions and confusion arising from the formative and summative uses of the same set of assessment evidence. Teachers need to be offered help on school-based assessment. This needs to be done in a way which disentangles the two different kinds of assessment so as to enable teachers to use assessment in a genuinely formative way to help students' learning. This would include guidance on types of feedback from teachers necessary to maintain pupil motivation, as well as on identifying specific aspects of attainment or good performance and what to do to help further improvement (see, for example, Tunstall and Gipps 1996). All these skills and understandings need to be developed within initial teacher training and continuing professional development.

Providing teachers with concrete examples illustrating how the problems can be tackled is certainly of help. However, this alone does not suffice to prepare teachers for the wide range of possible situations in school-based assessment where the contextual variables among schools vary so much. Teachers *themselves* (with appropriate professional training) need to analytically discern comparatively abstract, generative aspects of concrete assessment situations and relate them to each other (like what the present paper is attempting to do). According to Marton (1994), this is the way in which means-ends relationship or relationship between what is specific and what is general are established. He claims that for a teacher, "the transcendence of his taken for granted experiential world means the ascent of a kind of analytic awareness: a capability of abstracting aspects of concrete situations and seeing these aspects of concrete situations in relation to each other" (p.39). He asserts that such knowledge is theoretical knowledge in the sense that it is generalizable and generative, it is, or it can be decontextualizable and made explicit. And that analytic awareness is a function of theoretical knowledge, the availability of theoretical knowledge.

In the context of the present study, this theoretical knowledge is an understanding of the formative and summative functions of assessment as viewed from the educational assessment paradigm (Gipps 1994). In this respect, Harlen and James (1997) have offered many suggestions to help teachers to cope with the problems and to ease their tension. Their guiding principle (the

theoretical knowledge – in Marton's term) for the teachers is to *use* relevant evidence gathered as part of teaching for formative purposes but to *review* it, for summative purposes, in relation to the criteria which will be used for all students. This means that in the course of the practical teachers could still help individual students with their specific learning difficulties as far as they can review all these evidence, for a summative purpose, in relation to a set of criteria that are common to all students. Indeed, the TAS does not bar teachers from offering help to their students as long as they take into consideration of the help offered when giving the final assessment marks.

One final point to make is that the case of teacher A points to the need of assisting teachers to identify the referents - such as personal beliefs and myths - that they use to make sense of assessment practices. Tobin and Jakubowski (1990) assert that this is an important first step in beginning the process of change. Indeed, the findings of this study concur with that of Briscoe's study (1993) that the reconstruction process is particularly difficult when the referents a teacher uses to assign meaning to salient assessment roles and practices are cultural myths or beliefs associated with the myths. Accordingly, the impact of the culture of teaching on the construction and reconstruction of cognitive referents must be considered as teachers are assisted to initiate alternative assessment practices – like those advocated by the new forms of school-based assessment.

Conclusions

Overall, the findings of this study indicates how complex the assessment process is in the school laboratory, and point to the need for teachers to be aware of its potential for the improvement of teaching and learning and of its side-effects, especially when it is not properly carried out. There is clearly a need for teacher professional development in this vital aspect, especially on developing teachers' analytic awareness on the issue of fairness. It should be undertaken with full regard to problems of teacher interpretation and mediation at the classroom level. Special attention has to be paid on how to get the philosophy and intentions behind the new form of assessment across to the teachers - be it through initial or in-service teacher training. Otherwise, there is a danger of teachers interpreting these changes within a traditional 'testing paradigm' as in the case of teacher A and this will ruin the best intentions behind these new forms of school-based assessment. And this would be grossly unfair to all parties concerned – teachers and students alike.

References

Benson, G. D. (1989) The misrepresentation of science by philosophers and teachers of science. *Synthese*, 80, 107-119.

Biggs, J. (1996) The assessment scene in Hong Kong. In J. Biggs (ed), *Testing: To Educate or to Select? : Education in Hong Kong at the Crossroad*, (Hong Kong: Hong Kong Educational Publishing), 3-12.

Brickhouse, N. and Bodner, G. (1992) The beginning science teacher: Classroom narratives of convictions and constraints. *Journal of Research in Science Teaching*, 29(5), 471-485.

Briscoe, C. (1993) Using cognitive referents in making sense of teaching: A chemistry teacher's struggle to change assessment practices, *Journal of Research in Science Teaching*, 30, 971-987.

Buchan, A. S. (1993) Policy into practice: The operation of practical assessment in GCSE. *School Science Review*, 75, 7-15.

Cronin-Jones, L. and Shaw, S. (1992) The influence on the beliefs of preservice elementary and secondary teachers: Preliminary comparative analyses. *School Science and Mathematics*, 92, 14-22.

Donnelly, J. F., Buchan, A. S., Jenkins, E. W. & Welford, A. G. (1994) *Policy, Practice and Professional Judgement: School-based Assessment of Practical Science* (East Yorkshire: Studies in Education).

Duschl, R. A. & Wright, E. (1989) A case study of high school teachers' decision models for planning and teaching science. *Journal of Research in Science Teaching*, 26, 467-502.

Erickson, F. (1986) Qualitative methods in research on teaching. In M. C. Wittrock (ed), *Handbook of Research on Teaching* (New York: Macmillan), 119-161.

Fischler, H. (1994) Concerning the difference between intention and action: Teachers' conceptions and actions in physics teaching. In I. Carlgren, G. Handal, and S. Vaage (eds), *Teachers' minds and actions* (London: Falmer Press), 165-180.

Gipps, C. V. (1994) *Beyond Testing: Towards a Theory of Educational Assessment* (London: Falmer).

Guba, E. and Lincoln, Y. S. (1989) *Fourth Generation Evaluation* (Beverly Hills: Sage).

Hargreaves, A. (1989) *Curriculum and Assessment Reform* (Milton Keynes: Open University Press).

Harlen, W. (1994) Issues and approaches to quality assurance and quality control in assessment. In W. Harlen (ed), *Enhancing quality in assessment* (London: BERA), 11-25.

Harlen, W., Gipps, C., Broadfoot, P. and Nutall, D. (1992) Assessment and the improvement of education. *The Curriculum Journal*, 3(3).

Harlen, W. and James, M. (1997) Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education*, 4(3), 365-379.

Hewson, P. W. & Hewson, M. G. A' B. (1989) Analysis and Use of a Task for Identifying Conceptions of Teaching Science. *Journal of Education for Teaching*, 15(3), 191-209.

Hodson, D. (1992) Assessment of practical work: some considerations in philosophy of science. *Science and Education*, 1, 115-144.

Kagan, D. and Tappins, D. (1991) How teachers' classroom cases express their pedagogical beliefs. *Journal of Teachers Education*, 281-291.

Lakin, S. and Wellington, J. (1994) Who will teach the "nature of science"? : Teachers' views of science and their implications for science education. *International Journal of Science Education*, 16(2), 175-190.

Lantz, O. and Kass, H. (1987) Chemistry teachers' functional paradigms. *Science Education*, 71(1), 117-134.

Marton, F. (1994) On the structure of teachers' awareness. In I. Carlgren, G. Handal, and S. Vaage (eds), *Teachers' minds and actions* (London: Falmer Press), 29-42.

Marton, F. and Booth, S. (1997) *Learning and Awareness* (New Jersey: Lawrence Erlbaum Associaties).

McRobbie, C. J. and Tobin, K. (1995) Restrains to reform: The congruence of teacher and student actions in a chemistry classroom. *Journal of Research in Science Education*, 32(4), 373-385.

Mitchener and Anderson (1989) Teachers' perspective: developing and implementing an STS curriculum. *Journal of Research in Science Education*, 26(4), 351-369.

Morris, P. (1985) Teachers' perceptions of the barriers to the implementation of a pedagogic innovation: A South East Asian case study. *International Review of Education*, 31, 3-18.

Nyiri, J. C. (1988) Tradition and practical knowledge. In J. C. Nyiri and B. Smith (eds) *Practical Knowledge: Outline of a Theory of Traditions and Skills* (London: Croom Helm), 17-52.

Pang, K. C. (1992) The biology teacher assessment scheme (TAS). *Curriculum Forum*, 2 (2), 81-90.

Ropes (1998) *Review of Public Examination System in Hong Hong*. Interim Report (Hong Kong: Hong Kong Examinations Authority, Hong Kong Baptist University).

Sadler, R. (1989) Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

Sadler, R. (1998) Formative assessment: revisiting the territory. *Assessment in Education*, 5(1). 77-84.

Thompson, J. J. (1975) *Practical Work in Sixth Form Science* (Oxford: Department of Educational Studies, University of Oxford).

Tobin, K. and Jakubowski, E. (1990) *Cognitive processes and teacher change.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Tobin, K., Kahle, J. and Fraser, B. (1990) *Windows into the Science Classroom: Problem Associated with High Level Cognitive Learning in Science* (Philadelphia: Falmer Press).

Tobin, K. and LaMaster, S. (1995) Relationships between metaphors, beliefs and actions in a context of science curriculum change. *Journal of Research in Science Teaching*, 32(3), 225-242.

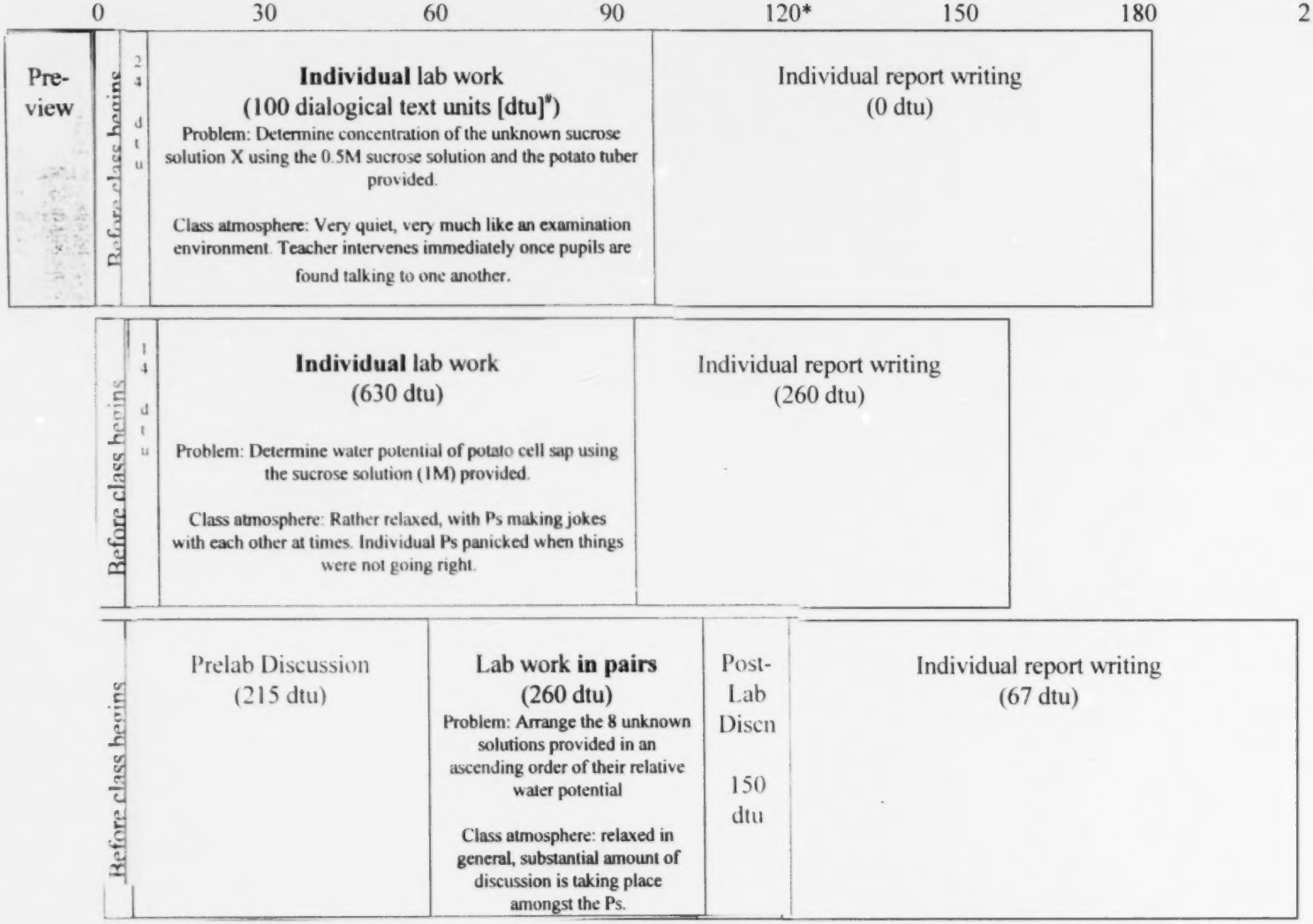
Tobin, K. and McRobbie, C. (1996) Cultural Myths as Constraints to the Enacted Science Curriculum. *Science Education*, 80(2), 223-241.

Torrance, H. (1995) Teacher involvement in new approaches to assessment. In H. Torrance (ed), *Evaluating Authentic Assessment* (Buckingham: OUP).

Tunstall, P. & Gipps, C. (1996) Teacher feedback to young children in formative assessment: a typology. *British Educational Research Journal*, 22, 389-404.

Yung, B. H. W. (1992) *An evaluation study of the feasibility of a teacher assessment scheme in the Hong Kong Advanced Level Biology Examination*. Unpublished M.Ed. thesis (Hong Kong: The University of Hong Kong).

Yung, B. H. W. (1995) Fostering professional competence and development: Experience of the HKAL Biology Teacher Assessment Scheme. In F. Lopez-Real (ed), *Teacher Education in the Asian Region. Proceedings of ITEC'95*. (Hong Kong: University of Hong Kong), 400-408.


0
30
60
90
120*
150
180
210 (min)
Teacher A
1. Female
2. 47 yrs old
3. 16 yrs AL teaching experience
4. 22 Ps, girls
Pre-view
Before class begins
24 dtu
Individual lab work
(100 dialogical text units [dtu]#)
Problem: Determine concentration of the unknown sucrose solution X using the 0.5M sucrose solution and the potato tuber provided.
Class atmosphere: Very quiet, very much like an examination environment. Teacher intervenes immediately once pupils are found talking to one another.
Individual report writing
(0 dtu)
Teacher B
1. Male
2. 28 yrs old
3. 6 yrs AL teaching experience
4. 8 Ps, co-ed
Before class begins
14 dtu
Individual lab work
(630 dtu)
Problem: Determine water potential of potato cell sap using the sucrose solution (1M) provided.
Class atmosphere: Rather relaxed, with Ps making jokes with each other at times. Individual Ps panicked when things were not going right.
Individual report writing
(260 dtu)
Teacher C
1. Female
2. 44 yrs old
3 18 yrs AL teaching experience
4 17 Ps, co-ed
Before class begins
Prelab Discussion
(215 dtu)
Lab work in pairs
(260 dtu)
Problem: Arrange the 8 unknown solutions provided in an ascending order of their relative water potential
Class atmosphere: relaxed in general, substantial amount of discussion is taking place amongst the Ps.
Post-Lab Discn
150 dtu
Individual report writing
(67 dtu)

N.B. * The normal lesson time for all the three cases is 120 minutes.

A dialogical text unit [dtu] is a unit of dialogue exchange between teacher and pupils.

Figure 1: Organization and Interaction Patterns of a Practical Work on Water Potential of Potato Cell Sap

BEST COPY AVAILABLE